



OICE

RAW SEQUENCE LISTING

DATE: 01/28/2002

PATENT APPLICATION: US/09/988,974

TIME: 11:43:25

Input Set : N:\Crf3\RULE60\09988974.raw

Output Set: N:\CRF3\01282002\I988974.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Hillman, Jennifer L.

6 Goli, Surya K.

7 Bandman, Olga

9 (ii) TITLE OF INVENTION: NOVEL RAB PROTEINS

12 (iii) NUMBER OF SEQUENCES: 9

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

16 (B) STREET: 3174 Porter Drive

17 (C) CITY: Palo Alto

18 (D) STATE: CA

19 (E) COUNTRY: USA

20 (F) ZIP: 94304

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette

24 (B) COMPUTER: IBM Compatible

25 (C) OPERATING SYSTEM: DOS

26 (D) SOFTWARE: FastSEQ for Windows Version 2.0

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/988,974

C--> 30 (B) FILING DATE: 19-Nov-2001

31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 09/215,887

35 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Billings, Lucy J.

39 (B) REGISTRATION NUMBER: 36,749

40 (C) REFERENCE/DOCKET NUMBER: PF-0183 US

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 415-855-0555

44 (B) TELEFAX: 415-845-4166

45 (C) TELEX:

48 (2) INFORMATION FOR SEQ ID NO: 1:

50 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 222 amino acids

52 (B) TYPE: amino acid

53 (C) STRANDEDNESS: single

54 (D) TOPOLOGY: linear

56 (vii) IMMEDIATE SOURCE:

57 (A) LIBRARY: Consensus

58 (B) CLONE: Consensus

ENTERED

RAW SEQUENCE LISTING

DATE: 01/28/2002

PATENT APPLICATION: US/09/988,974

TIME: 11:43:25

Input Set : N:\Crif3\RULE60\09988974.raw

Output Set: N:\CRF3\01282002\I988974.raw

```

60      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
62 Met Ser Asp Ser Glu Glu Ser Gln Asp Arg Gln Leu Lys Ile Val
63 1          5          10          15
W--> 64 Val Leu Gly Asp Xaa Ala Ser Gly Lys Thr Ser Leu Thr Thr Cys Phe
65          20          25          30
66 Ala Gln Glu Thr Phe Gly Lys Gln Tyr Lys Gln Thr Ile Gly Leu Asp
67          35          40          45
68 Phe Phe Leu Arg Arg Ile Thr Leu Pro Gly Asn Leu Asn Val Thr Leu
69          50          55          60
70 Gln Ile Trp Asp Ile Gly Gly Gln Thr Ile Gly Gly Lys Met Leu Asp
71 65          70          75          80
72 Lys Tyr Ile Tyr Gly Ala Gln Gly Val Leu Leu Val Tyr Asp Ile Thr
73          85          90          95
74 Asn Tyr Gln Ser Phe Glu Asn Leu Glu Asp Trp Tyr Thr Val Val Lys
75          100         105         110
W--> 76 Lys Val Ser Xaa Glu Ser Glu Thr Gln Pro Leu Val Ala Leu Val Gly
77          115         120         125
78 Asn Lys Ile Asp Leu Glu His Met Arg Thr Ile Lys Pro Glu Lys His
79          130         135         140
80 Leu Arg Phe Cys Gln Glu Asn Gly Phe Ser Ser His Phe Val Ser Ala
81 145         150         155         160
82 Lys Thr Gly Asp Ser Val Phe Leu Cys Phe Gln Lys Val Ala Ala Glu
83          165         170         175
W--> 84 Ile Leu Gly Ile Lys Leu Asn Xaa Xaa Gln Xaa Xaa Xaa Ser His Xaa
85          180         185         190
W--> 86 Gly Val Val Lys Xaa Xaa Ile Val Asn Tyr Asn Gln Glu Pro Met Ser
87          195         200         205
W--> 88 Arg Thr Xaa Asn Pro Pro Arg Ser Ser Met Cys Ala Val Gln
89          210         215         220

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 914 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

105 ATGGCGCTTG CCGAGTGATT CTCCTCGAAT ACCTCCTGCC GGCGCGGAGA CACCGGGGCG      60
106 GGGGTCTGTC CGCAACTACC TCCCTTCCTC CTCTCCCCCG CCCCCGGAGC CTTTCATCCTT      120
107 CCCTTCCCCC CCCACCTCGA GGGGCGGGCC TGGTTCCCGG GACACCATGT CGGACTCTGA      180
108 GGAGGAGAGC CAGGACCGGC AACTGAAAT CGTCGTGCTG GGGGACGNGG CCTCCGGGAA      240
109 GACCTCCTTA ACTACGTGTT TTGCTCAAGA AACTTTTGGG AAACAGTACA AACAAACTAT      300
110 AGGACTGGAT TTCTTTTTGA GAAGGATAAC ATTGCCAGGA AACTTGAATG TTACCCTTCA      360
111 AATTTGGGAT ATAGGAGGGC AGACAATAGG AGGCAAAATG TTGGATAAAT ATATCTATGG      420
112 AGCACAGGGA GTCCTCTTGG TATATGATAT TACAAATTAT CAAAGCTTTG AGAATTTAGA      480
113 AGATTGGTAT ACTGTGGTGA AGAAAGTGAG CNAGGAGTCA GAAACTCAGC CACTGGTTGC      540
114 CTTGGTAGGC AATAAAATTG ATTTGGAGCA TATGCGAACA ATAAACCTG AAAAACACTT      600

```

RAW SEQUENCE LISTING

DATE: 01/28/2002

PATENT APPLICATION: US/09/988,974

TIME: 11:43:25

Input Set : N:\Crf3\RULE60\09988974.raw

Output Set: N:\CRF3\01282002\I988974.raw

```

115 ACGGTTTTTGC CAGGAAATG GTTTTAGTAG CCACTTTGTC TCAGCCAAGA CAGGAGACTC 660
116 TGTCTTCCTG TGCTTTCAGA AAGTTGCTGC TGAAATCCTT GGGATCAAAT TAAACAANNA 720
117 GCAGAAWTRG MACAGTCACA GWGGGGTGGT GAAGGSAGRT ATTGTAAACT ACAACCAGGA 780
118 ACCTATGTCA AGGACTKTTA ACCCTCCTAG AAGCTCTATG TGTGCAGTTC AGTGAGCGCA 840
119 TTTTNCITTTT GTNTTGATAG TTCTGGCTGC CCTTCAACTC TGGGTGGGNC CCNAGGGCTT 900
120 CTAGGACTTG TTTT 914

```

122 (2) INFORMATION FOR SEQ ID NO: 3:

124 (i) SEQUENCE CHARACTERISTICS:

125 (A) LENGTH: 217 amino acids

126 (B) TYPE: amino acid

127 (C) STRANDEDNESS: single

128 (D) TOPOLOGY: linear

130 (vii) IMMEDIATE SOURCE:

131 (A) LIBRARY: Consensus

132 (B) CLONE: Consensus

134 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

W--> 136 Met Arg Xaa Lys Met Gly Asn Gly Thr Glu Glu Asp Tyr Asn Phe Val
137 1 5 10 15
138 Phe Lys Val Val Leu Ile Gly Glu Ser Gly Val Gly Lys Thr Asn Leu
139 20 25 30
140 Leu Ser Arg Phe Thr Arg Asn Glu Phe Ser His Asp Ser Arg Thr Thr
141 35 40 45
142 Ile Gly Val Glu Phe Ser Thr Arg Thr Val Met Leu Gly Thr Ala Ala
143 50 55 60
144 Val Lys Ala Gln Ile Trp Asp Thr Ala Gly Leu Glu Arg Tyr Arg Ala
145 65 70 75 80
146 Ile Thr Ser Ala Tyr Tyr Arg Gly Ala Val Gly Ala Leu Leu Val Phe
147 85 90 95
148 Asp Leu Thr Lys His Gln Thr Tyr Ala Val Val Glu Arg Trp Leu Lys
149 100 105 110
150 Glu Leu Tyr Asp His Ala Glu Ala Thr Ile Val Val Met Leu Val Gly
151 115 120 125
152 Asn Lys Ser Asp Leu Ser Gln Gly Arg Glu Val Pro Thr Glu Glu Ala
153 130 135 140
154 Arg Met Phe Ala Glu Asn Asn Gly Leu Leu Phe Leu Glu Thr Ser Ala
155 145 150 155 160
156 Leu Asp Ser Thr Asn Val Glu Leu Ala Phe Glu Thr Val Leu Lys Glu
157 165 170 175
158 Ile Phe Ala Lys Val Ser Lys Gln Arg Gln Asn Ser Ile Arg Thr Asn
159 180 185 190
W--> 160 Ala Ile Thr Leu Gly Ser Ala Gln Xaa Gly Gln Glu Pro Gly Pro Gly
161 195 200 205
162 Glu Lys Arg Ala Cys Cys Ile Ser Leu
163 210 215

```

165 (2) INFORMATION FOR SEQ ID NO: 4:

167 (i) SEQUENCE CHARACTERISTICS:

168 (A) LENGTH: 847 base pairs

169 (B) TYPE: nucleic acid

170 (C) STRANDEDNESS: single

RAW SEQUENCE LISTING

DATE: 01/28/2002

PATENT APPLICATION: US/09/988,974

TIME: 11:43:25

Input Set : N:\Crf3\RULE60\09988974.raw

Output Set: N:\CRF3\01282002\I988974.raw

```

171         (D) TOPOLOGY: linear
173     (vii) IMMEDIATE SOURCE:
174         (A) LIBRARY: Consensus
175         (B) CLONE: Consensus
177     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
179 GGCATTGAGC CAACACACAG ATTTGTCGCC TCTGTCCCCG AAGACACCTG CACCCTCCAT      60
180 GCGGANCAAG ATGGGGAATG GAACTGAGGA AGATTATAAC TTTGTCTTCA AGGTGGTGCT      120
181 GATCGGCGAA TCAGGTGTGG GGAAGACCAA TCTACTCTCC CGATTACGC GCAATGAGTT      180
182 CAGCCACGAC AGCCGCACCA CCATCGGGGT TGAGTTCTCC ACCCGCACTG TGATGTTGGG      240
183 CACCGCTGCT GTCAAGGCTC AGATCTGGGA CACAGCTGGC CTGGAGCGGT ACCGAGCCAT      300
184 CACCTCGGCG TACTATCGTG GTGCAGTGGG GGCCCTCTCG GTGTTTGACC TAACCAAGCA      360
185 CCAGACCTAT GCTGTGGTGG AGCGATGGCT GAAGGAGCTC TATGACCATG CTGAAGCCAC      420
186 GATCGTCGTC ATGCTCGTGG GTAACAAAAG TGACCTCAGC CAGGGCCGGG AAGTGCCAC      480
187 TGAGGAGGCC CGAATGTTCG CTGAAAACAA TGGACTGCTC TTCCTGGAGA CCTCAGCCCT      540
188 GGACTCTACC AATGTTGAGC TAGCCTTTGA GACTGTCTCG AAAGAAATCT TTGCGAAGGT      600
189 GTCCAAGCAG AGACAGAACA GCATCCGGAC CAATGCCATC ACTCTGGGCA GTGCCCAGGN      660
190 TGGACAGGAG CCTGGCCCTG GGGAGAAGAG GGCCTGTTGC ATCAGCCTCT GACCTTGGCC      720
191 AGCACCACCT GCCCCCACTG GCTTTTGGT GCCCCTTGTC CCCACTTCAG CCCAGGACC      780
192 TTTCTTGCC CTTTGGTTCC AGATATCAGA CTGTTCCCTG TTCAcAGCAC CCTCAGGGTC      840
193 TTAAGGT
195 (2) INFORMATION FOR SEQ ID NO: 5:
197     (i) SEQUENCE CHARACTERISTICS:
198         (A) LENGTH: 201 amino acids
199         (B) TYPE: amino acid
200         (C) STRANDEDNESS: single
201         (D) TOPOLOGY: linear
203     (vii) IMMEDIATE SOURCE:
204         (A) LIBRARY: Consensus
205         (B) CLONE: Consensus
207     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
209 Met Ala Gly Lys Ser Ser Leu Phe Lys Val Ile Leu Leu Gly Asp Gly
210 1          5          10          15
211 Gly Val Gly Lys Ser Ser Leu Met Asn Arg Tyr Val Thr Asn Lys Phe
212          20          25          30
213 Asp Thr Gln Leu Phe His Thr Ile Gly Val Glu Phe Leu Asn Lys Asp
214          35          40          45
215 Leu Glu Val Asp Gly His Phe Val Thr Met Gln Ile Trp Asp Thr Ala
216          50          55          60
217 Gly Gln Glu Arg Phe Arg Ser Leu Arg Thr Pro Phe Tyr Arg Gly Ser
218          65          70          75          80
219 Asp Cys Cys Leu Leu Thr Phe Ser Val Asp Asp Ser Gln Ser Phe Gln
220          85          90          95
221 Asn Leu Ser Asn Trp Lys Lys Glu Phe Ile Tyr Tyr Ala Asp Val Lys
222          100         105         110
223 Glu Pro Glu Ser Phe Pro Phe Val Ile Leu Gly Asn Lys Ile Asp Ile
224          115         120         125
225 Ser Glu Arg Gln Val Ser Thr Glu Glu Ala Gln Ala Trp Cys Arg Asp
226          130         135         140
227 Asn Gly Asp Tyr Pro Tyr Phe Glu Thr Ser Ala Lys Asp Ala Thr Asn

```

RAW SEQUENCE LISTING

DATE: 01/28/2002

PATENT APPLICATION: US/09/988,974

TIME: 11:43:25

Input Set : N:\Crf3\RULE60\09988974.raw

Output Set: N:\CRF3\01282002\I988974.raw

```

228 145          150          155          160
229 Val Ala Ala Ala Phe Glu Glu Ala Val Arg Arg Val Leu Ala Thr Glu
230          165          170          175
231 Asp Arg Ser Asp His Leu Ile Gln Thr Asp Thr Val Asn Leu His Arg
232          180          185          190
233 Lys Pro Lys Pro Ser Ser Ser Cys Cys
234          195          200
236 (2) INFORMATION FOR SEQ ID NO: 6:
238 (i) SEQUENCE CHARACTERISTICS:
239 (A) LENGTH: 1175 base pairs
240 (B) TYPE: nucleic acid
241 (C) STRANDEDNESS: single
242 (D) TOPOLOGY: linear
244 (vii) IMMEDIATE SOURCE:
245 (A) LIBRARY: Consensus
246 (B) CLONE: Consensus
248 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
250 CACTGTGATG AAACACTTTT CCCGTGTCGT TTGAGTGCAT CTTCTCAACA ACCCTAGGAG 60
251 GGTTCTTGAA GCTTTTGAGA TTAACAATGG CAGGAAAATC ATCACTTTTT AAAGTAATTC 120
252 TCCTTGGAGA TGGTGGAGTT GGGAAGAGTT CACTATGAA CAGATATGTA ACTAATAAGT 180
253 TTGATACCCA GCTCTTCCAT ACAATAGGTG TGGAATTTTT AAATAAAGAT TTGGAAGTGG 240
254 ATGGACATTT TGTTACCATG CAGATTTGGG ACACGGCAGG TCAGGAGCGA TTCCGAAGCC 300
255 TGAGGACACC ATTTTACAGA GGTTCGTACT GCTGCCTGCT TACTTTTAGT GTCGATGATT 360
256 CACAAAGCTT CCAGAACTTA AGTAACTGGA AGAAAGAATT CATATATTAT GCAGATGTGA 420
257 AAGAGCCTGA GAGCTTTCCT TTTGTGATTC TGGGTAACAA GATTGACATA AGCGAACGGC 480
258 AGGTGTCTAC AGAAGAAGCC CAAGCTTGGT GCAGGGACAA CGGCGACTAT CCTTATTTTG 540
259 AAACAAGTGC AAAAGATGCC ACAAATGTGG CAGCAGCCTT TGAGGAAGCG GTTCGAAGAG 600
260 TTCTTGCTAC CGAGGATAGG TCAGATCATT TGATTGAGAC AGACACAGTC AATCTTCACC 660
261 GAAAGCCCAA GCCTAGCTCA TCTTGCTGTT GATTGTTAGA TTGTTGATGC ATTCTAACCA 720
262 ACTCACACAT ATACACAAAA TCAACATGGG GATGGAGAAG AGAATTAGCG TTTGCAGCAG 780
263 TGTATCATCT ACTAATAAAA TTAAACTAAT GTTGCTGCTT CATTAGTTGG TGGGAGAAGG 840
264 GACACATCCA CTCTTGAGG AATATATTTA CTCAATAATG GCACCTTACA TTTATAAATT 900
265 GTAACAGTTG TCTAATAACG TTTCTTTAAT TTAAATATGT AAGTTGCAGA GCTAATAAAT 960
266 GAAATGACCA AGACTTTAAT TATAATAAAA ATAAGAACT TGACTATTCT AGAAGTTATA 1020
267 CTTGGATTTT TTCCTGGGAA AATGGAGAAC TACTTTTTAT ATGTGTATGT TTTTATGCAA 1080
268 TTAGCATTGT ATTCTTGGTT CAGGGAAATA CTTTCCTAAA GCAATAATGT TAGATATTAA 1140
269 AGATTAAAAT CTAATGTAAA AAAAAAAAAA AAAAA 1175
271 (2) INFORMATION FOR SEQ ID NO: 7:
273 (i) SEQUENCE CHARACTERISTICS:
274 (A) LENGTH: 221 amino acids
275 (B) TYPE: amino acid
276 (C) STRANDEDNESS: single
277 (D) TOPOLOGY: linear
279 (vii) IMMEDIATE SOURCE:
280 (A) LIBRARY: GenBank
281 (B) CLONE: 1154901
283 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
285 Met Ser Asp Ser Glu Glu Glu Ser Gln Asp Arg Gln Leu Lys Ile Val
286 1 5 10 15

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/988,974

DATE: 01/28/2002

TIME: 11:43:26

Input Set : N:\Crf3\RULE60\09988974.raw

Output Set: N:\CRF3\01282002\I988974.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:64 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:76 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:84 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:86 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:88 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:136 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:160 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3